Algorithm used: Gradient Decent

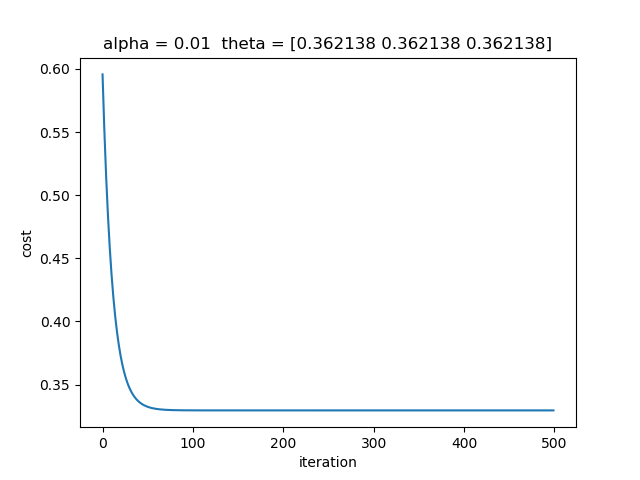
Normalized data through taking the difference of X from the mean and then divided by the standard deviation

**Round 1**

MAX\_iter = 500

Alpha = .01

results: 0.6434164769124346 (0.5006554500688641)



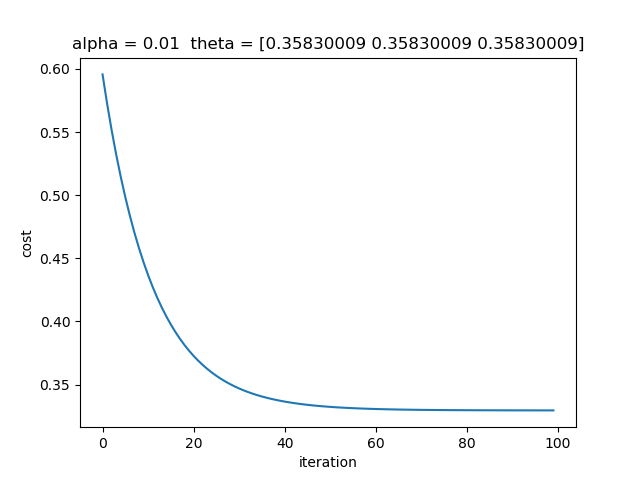
I noticed that it didn’t take 500 iterations to converge so, I changed the MAX\_ITER to 100, not changing the learning rate

**Round 2**

MAX\_iter = 100

Alpha = .01

results: 0.6381631832667771 (0.5006863528964408)



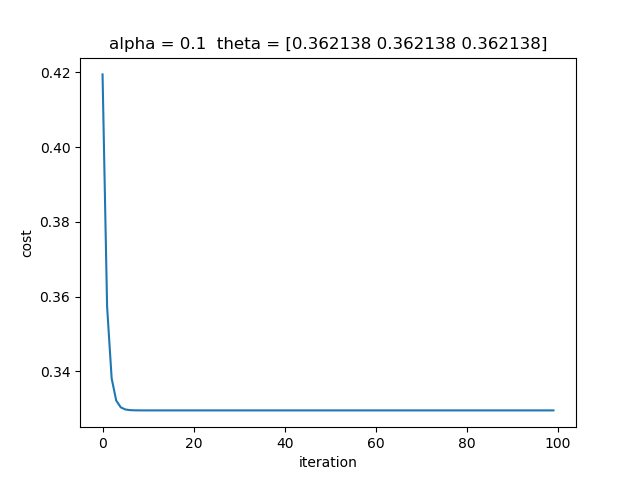
Notice results were still pretty accurate

**Round 3**

MAX\_iter = 100

Alpha = .1

results: 0.6434164769787035 (0.5006554500690491)



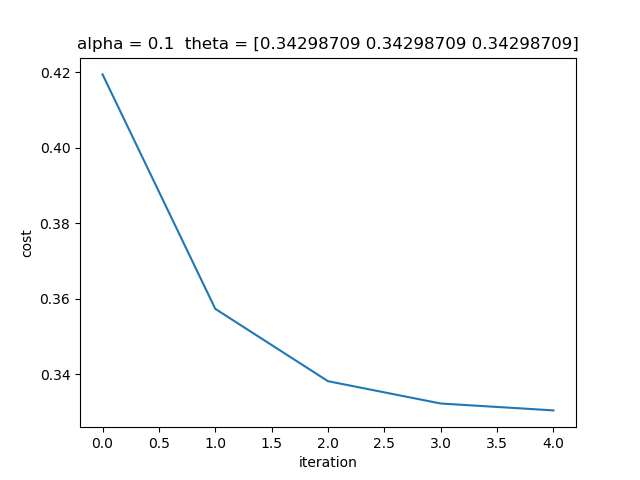
If I change theta to be a bigger number the gradient decent converges faster, but the results are still similar

**Round 4**

MAX\_iter = 4

Alpha = .1

results: 0.6176752141316211 (0.5011339226254662)



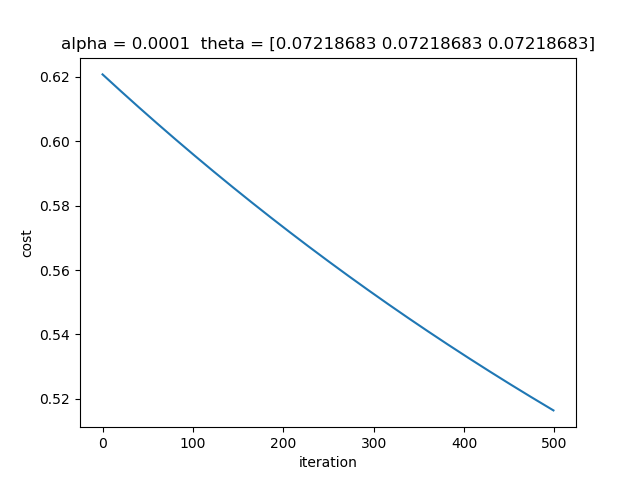
Notice that the results output changed

**Round 4**

MAX\_iter = 500

Alpha = .0001

results: 0.5310077119459345 (0.5373411740826544)



I chose a way smaller learning rate, and noticed that the convergence happened over a larger amount of iterations, but the results seem to be more accurate and is a stricter model